

Machine learning approaches for wind speed forecasting using long-term monitoring data: a comparative study

X.W. Ye^a, Y. Ding^b and H.P. Wan*

Department of Civil Engineering, Zhejiang University, Hangzhou 310058, China

(Received May 26, 2018, Revised July 30, 2019, Accepted August 6, 2019)

Abstract. Wind speed forecasting is critical for a variety of engineering tasks, such as wind energy harvesting, scheduling of a wind power system, and dynamic control of structures (e.g., wind turbine, bridge, and building). Wind speed, which has characteristics of random, nonlinear and uncertainty, is difficult to forecast. Nowadays, machine learning approaches (generalized regression neural network (GRNN), back propagation neural network (BPNN), and extreme learning machine (ELM)) are widely used for wind speed forecasting. In this study, two schemes are proposed to improve the forecasting performance of machine learning approaches. One is that optimization algorithms, i.e., cross validation (CV), genetic algorithm (GA), and particle swarm optimization (PSO), are used to automatically find the optimal model parameters. The other is that the combination of different machine learning methods is proposed by finite mixture (FM) method. Specifically, CV-GRNN, GA-BPNN, PSO-ELM belong to optimization algorithm-assisted machine learning approaches, and FM is a hybrid machine learning approach consisting of GRNN, BPNN, and ELM. The effectiveness of these machine learning methods in wind speed forecasting are fully investigated by one-year field monitoring data, and their performance is comprehensively compared.

Keywords: structural health monitoring; wind speed prediction; machine learning; optimization algorithm; finite mixture method

1. Introduction

It is essential to forecast wind speed for a variety of engineering tasks, including but not limited to the integration of wind energy into electricity grids, scheduling of a wind power system, and design of reliable structures (e.g., wind turbine, bridge, and building). Wind effects can cause destructive damage to infrastructure systems and significant human and economic losses (Ye *et al.* 2015a). It was reported that if the wind loading exceeds a certain value, the high-rise buildings and the long-span bridges, which have the characteristics of light mass, high flexibility and slight damping, may be seriously affected (Ye *et al.* 2016a, Chen *et al.* 2014, Ye *et al.* 2012, Ni *et al.* 2012, Ni *et al.* 2010). To ensure reliable wind-resistant design, it is important to forecast wind speed to precisely assess the wind-induced effects in advance (Ye *et al.* 2017, Ye *et al.* 2016b, Hocaoglu *et al.* 2007).

Accurate forecasting of wind speed is challenging since it usually has the characteristics of random, nonlinear, and uncertainty. In the literature, a lot of methods have been explored to forecast wind speed. Physical models, conventional statistical models, and machine learning methods are commonly used for the wind speed forecasting (Lei *et al.* 2009). For the physical model, it uses physical or

meteorological information such as the temperature, pressure, and orography to evaluate the future speed, which depends on the known samples and fails to have long-term prediction (Landberg 1999). For the conventional statistical model, it uses historical wind speed data for training and its goal is to find the relationship between certain explanatory variable and future wind speed, which is difficult to develop (Khalid *et al.* 2012). For the machine learning method, it just uses time-series data (called training data) and is powerful to simulate complex nonlinear systems (Li *et al.* 2009). Since more and more structures have installed or will install structural health monitoring (SHM) systems, the field measurements, which considers the real geographical environment, terrain roughness, structural shape, etc., provide reliable training data for machine learning methods (Ye *et al.* 2016c, Ye *et al.* 2015b, Ye *et al.* 2013). In this regard, the machine learning methods are being widely used in forecasting the wind speed.

A lot of machine learning approaches have been developed to predict the wind speed, such as generalized regression neural network (GRNN), back propagation neural network (BPNN), extreme learning machine (ELM). Kumar *et al.* (2016) introduced GRNN for long-term wind speed prediction of major wind power potential states in Western Region of India. Lee *et al.* (2012) adopted the GRNN to predict wind speed obtained from Chiang Kai-shek International Airport. Huang *et al.* (2017) utilized BPNN to predict the time series of wind-induced pressures on a building surface. Guo *et al.* (2011) established a wind speed forecasting method based on a BPNN. Liu *et al.* (2018) used ELM method to predict the wind speed data decomposed by wavelet packet and empirical mode method.

*Corresponding author, Ph.D., Professor
E-mail: hpwan@zju.edu.cn

^a Professor

^b Ph.D. Candidate

Lazarevska (2016) presented ELM method to model the wind speed on a short-term basis and the results show that the ELM method possesses the attributes of simplicity, good performance, and fast computation. The performance of the machine learning methods can be influenced by parameter setting, such as weight, threshold, learning rate. The parameter setting determined by people's experience may undermine the prediction performance (Huang *et al.* 2010). In fact, most of the parameters involved in machine learning approaches can be automatically determined by effective optimization algorithms.

A variety of optimization algorithms could be employed to improve the accuracy of the machine learning methods, such as cross validation (CV), genetic algorithm (GA), particle swarm optimization (PSO). Wang *et al.* (2018) determined the optimal spread factor which is the key parameter of the GRNN by CV method. Jiang *et al.* (2016) proposed a GRNN with K -fold CV method for predicting the displacement of landslide. Kassa *et al.* (2016) used the GA to optimize the weights and biases of BPNN by the modeling dataset. Xu *et al.* (2011) presented a wind power prediction model based on GA-BPNN which is shown to outperform the conventional BPNN. Han *et al.* (2013) used the PSO algorithm to determine the weights and thresholds of ELM and found that the PSO-ELM had better global approximation performance and generalization capability. Yang *et al.* (2014) proposed a modified ensemble of ELM based on PSO and it was shown that the modified method owned better convergence performance than some classical ELMs. In view of the above, the proper optimization algorithms can improve the accuracy of the machine learning methods in the wind forecasting. Since different machine learning methods have their own forecasting advantages, combination of different optimization algorithms should be considered. Finite mixture (FM) method is an effective tool to combine different methods (Ye *et al.* 2016d, McLachlan *et al.* 2000). Thus, the hybrid machine learning approach implemented by FM is also used for wind speed forecasting.

In this study, three machine learning approaches (GRNN, BPNN, and ELM) are adopted for wind speed forecasting. Furthermore, the optimization algorithm-assisted machine learning methods (CV-GRNN, GA-BPNN, and PSO-ELM) as well as the hybrid one (i.e., FM method) are developed for improving the forecasting performance. For the optimization algorithm-assisted approaches, the analytical fitness functions are derived. For the FM method, the analytical expressions for calculating the weights are derived. Finally, the effectiveness of these considered machine learning methods in wind speed forecasting are fully investigated by one-year field monitoring data, and their performance is comprehensively compared as well. The rest of paper is organized as follows. In Section 2, the GRNN, BPNN and ELM methods are described for wind speed forecasting. Section 3 details optimization algorithms, FM method as well as model assessment. In Section 4, the instrumented bridge is briefly introduced and wind filed measurements are demonstrated. The results of wind speed forecasting by the machine learning methods are given in Section 5. Finally, Section 6

ends with some conclusions drawn from this study.

2. Machine learning approaches

2.1 GRNN approach

Generalized regression neural network (GRNN) is a powerful regression tool with a dynamic network structure (Specht 1993). Structure of GRNN includes input layer, pattern layer, summation layer and output layer (Heimes *et al.* 1998). Generally, the GRNN is used for the estimation of continuous variables using nonlinear regression analysis (Cigizoglu 2005). Base on this definition, the dependent variable Y , which is network output, involves in an independent X , which is network input. The $f(x,y)$ represents the known joint continuous probability density function (PDF) of x and y . X is a particular measured value of the random variable x . The conditional mean of y given X can be expressed by (Specht 1991)

$$E(y | X) = \frac{\int_{-\infty}^{+\infty} yf(X, y) dy}{\int_{-\infty}^{+\infty} f(X, y) dy} \quad (1)$$

However, when the $f(x,y)$ is not known, it can be estimated from sample observations of X and Y (Specht 1991)

$$f(X, Y) = \frac{1}{(2\pi)^{(m+1)/2} \sigma^{m+1} n} \cdot \sum_{i=1}^n \exp \left\{ -\frac{(X - X_i)^T (X - X_i) + (Y - Y_i)^2}{2\sigma^2} \right\} \quad (2)$$

where X_i and Y_i is the sample observation of the random variables x and y ; m is the dimension of the random variable x ; n is the number of samples; and σ is the smoothing parameter. By simplifying the Eqs. (1) and (2), we have

$$E(y | X) = \frac{\sum_{i=1}^n Y_i \exp \left\{ -\frac{(X - X_i)^T (X - X_i)}{2\sigma^2} \right\}}{\sum_{i=1}^n \exp \left\{ -\frac{(X - X_i)^T (X - X_i)}{2\sigma^2} \right\}} \quad (3)$$

It can be known from Eq. (3) that the GRNN has only one smoothing parameter σ that needs to be determined, which is a core parameter in using GRNN for forecasting (Specht 1992). The prediction accuracy of GRNN is determined by the value of smoothing parameter σ . On the one hand, when the value of σ is larger, the estimated value of the $f(x,y)$ is smoother, which becomes a multivariate Gauss function. On the other hand, when the value of σ is smaller, the $f(x,y)$ becomes a non-Gauss model. Therefore, it is necessary to use an optimization algorithm to select best smoothing parameter σ such as cross validation (CV) (Li *et al.* 2013).

2.2 BPNN approach

Back propagation neural network (BPNN) is simply a gradient descent method designed to minimize the total error of the output computed by the network (Werbos 1974). In the structure of the BPNN, it consists of an input layer, an output layer, and one or more hidden layers between them. The hidden layer's upper limit can be determined base on the number of input layer and output layer, expressed by (Jadid *et al.* 1994)

$$P_{\text{hidden}} \leq n_{\text{train}}(n_{\text{input}} + n_{\text{output}}) / R \quad (4)$$

where P_{hidden} is the number of hidden layer; n_{input} is the number of input layer; n_{output} is the number of output layer; n_{train} is the number of sample training; and R is an adjustment constant, set as $5 \leq R \leq 10$.

The BPNN can express the function mapping relationship from n_{input} independent variables to n_{output} dependent variables, given by (Ding *et al.* 2011)

$$a_j = f_1\left(\sum_{i=1}^{n_{\text{input}}} w_{ij}x_i - \varphi_j\right) \quad (5)$$

$$t_{\text{kBP}} = f_2\left(\sum_{j=1}^{P_{\text{hidden}}} w_{jk}a_j - \varphi_k\right) \quad (6)$$

$$e_{\text{BP}} = \sum_{k=1}^{n_{\text{output}}} (t_{\text{kBP}} - y_{\text{kBP}})^2 \quad (7)$$

where x_i is the value of input; a_j is the output of the hidden layer; t_{kBP} is the output of the network; y_{kBP} is the desired output; e_{BP} is the error function; φ_j is the threshold value of the hidden layer; φ_k is the threshold value of the output layer; w_{ij} is the weight from input layer to hidden layer; w_{jk} is the weight from hidden layer to output layer; f_1 is the transfer function of the hidden layer; and f_2 is the transfer function of the output layer.

Once the error is obtained, it is input to the network and the weights and threshold have been continuously adjusted in the network, aiming to make the error be less than a pre-set minimum value. Then, input time instants are entered into the trained network to obtain the forecasting results. The initial weight and initial threshold are important parameters of the BPNN, which can be determined by optimization algorithms such as genetic algorithm (GA).

2.3 ELM approach

Extreme learning machine (ELM) is composed of input layers, hidden layers, and output layers. ELM is an efficient learning method for single-hidden layer feed-forward neural network (SLFN) (Huang *et al.* 2006). In the ELM, the input weights and thresholds can be determined implicitly by the hidden-layer output matrix, which can be expressed by (Huang *et al.* 2007)

$$f_{\text{ELM}}(x) = \sum_{l=1}^L h_l(x)\beta_l = \sum_{l=1}^L g(\omega_l x_i + b_l)\beta_l \quad (8)$$

where $f_{\text{ELM}}(x)$ is the output of the network; x_i is an input value; β_l is the output weights; L is the number of hidden layer; w_l is the input weights; b_l is the threshold; $h_l(x)$ is the

hidden layer function; and $g(\bullet)$ is an activation function.

The weight parameters β can be calculated by minimizing the objective loss function, written as

$$\varepsilon_{\text{ELM}} = \sum_{l=1}^L g(\omega_l x_i + b_l)\beta_l - y_{\text{ELM}} \quad (9)$$

$$\arg \min \|\varepsilon_{\text{ELM}}\|^2 = \|H\beta - Y\|^2 \quad (10)$$

where y_{ELM} is the expected output, $Y = \{y_{\text{ELM}1}, y_{\text{ELM}2}, \dots, y_{\text{ELM}L}\}$; ε_{ELM} is the error between the output value and the expected value; H is the hidden layer output matrix of the neural network, $H = \{h_1(x), h_2(x), \dots, h_L(x)\}$.

The minimization problem defined in the Eqs. (9) and (10) can be solved by well-known least square estimator

$$\beta = H^+ Y = H^T (H H^T)^{-1} Y \quad (11)$$

where H^+ is the Moore-Penrose generalized inverse of matrix H .

The performance of ELM largely relies on the given labels of training data and initial input weights and initial thresholds (Kai *et al.* 2016). For the training samples, the sample errors can be reduced by structural health monitoring (SHM) system, which can consider the real geographical environment, terrain roughness, structural shape, etc. Similarly, the input weights and thresholds can be determined by optimization algorithms such as particle swarm optimization (PSO).

3. Optimization algorithm-assisted and hybrid machine learning approaches

3.1 CV-GRNN approach

Cross validation (CV) is a measurement of assessing the performance of a predictive model, and performance assessment is done based on an independent dataset (Diana *et al.* 2002). Usually, CV is to divide data into two segments, which attempts to avoid over-fitting and obtain more valid information. One set is used to learn or train a model and the other set is used to validate the model (Refaeilzadeh *et al.* 2016). Many CV schemes are available, such as K -fold cross validation (K -CV) (Ping *et al.* 2014), leave-one-out cross validation (LOO-CV) (Shao *et al.* 2016). The LOO-CV algorithm needs a long computational time. Therefore, the K -CV algorithm is used to determine the best smoothing parameter of GRNN. Mean square error (MSE) is used as the fitness function of GRNN (Braganeto *et al.* 2004), given by

$$F_{\text{CVGRNN}} = \frac{1}{n} \sum_{i=1}^n (T_i - Y_i)^2 \quad (12)$$

where n is the number of samples; T_i is the network training output value; and Y_i is the expected output value; and F_{CVGRNN} is the fitness function.

The CV-GRNN model includes three important steps (Jiang *et al.* 2016): (i) the data is partitioned into K equally folds; (ii) a single sub-sample is regarded as the validation

data for testing the model and the remaining $K-1$ sub-samples are used as training data; and (iii) the optimal parameter are determined by fitness function.

3.2 GA-BPNN approach

Genetic algorithm (GA) is randomized search and optimization method, which guided by the principle of survival of the fittest (Holland *et al.* 1973). GA includes initial population, fitness function, and genetic operations (selection, crossover, and mutation). For the initial population, the parameters of the search space are encoded in the form of chromosomes. For the fitness function, the goodness of the chromosomes is determined. For the genetic operations, the process of selection, crossover and mutation continues for a fixed number of generations or till a termination condition is satisfied (Maulik *et al.* 2000). Therefore, the optimal parameters of BPNN can be determined by GA method. GA needs to establish a suitable fitness function (Cheung *et al.* 1997). The absolute value of the error between the predicted output and the expected output is used as the fitness function

$$F_{GABP} = \sum_{k=1}^m |t_{kBP} - y_{kBP}| \quad (13)$$

where m is the network output number; t_{kBP} is the output of the network; y_{kBP} is the expected output; and F_{GABP} is a fitness function.

GA-BPNN model consists of three important steps (Wang *et al.* 2016): (i) randomly generate initial weights and thresholds; (ii) modify the weights and thresholds by fitness function; and (iii) repeat calculation processes until the difference between the output of the network and the expected output is small enough.

3.3 PSO-ELM approach

Particle swarm optimization (PSO) is a global optimization algorithm in which the feasible solution can be represented as a point or surface in a multidimensional search space (Beheshti *et al.* 2014). The PSO algorithm searches the space of an objective function by adjusting the trajectories of individual agents, which is called particles (Engelbrecht 2006). Each particle has a memory function, and adjusts its trajectory according to two pieces of information, the best position that it has so far visited, and the global best position attained by the whole swarm (Kennedy *et al.* 1995). The velocity and displacement parameters of each particle in the swarm are updated at each iteration, expressed by (Wen *et al.* 2010)

$$v_i(t+1) = \omega v_i(t) + c_1 r_1 (p_i(t) - x_i(t)) + c_2 r_2 (g_i(t) - x_i(t)) \quad (14)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (15)$$

$$F_{PSOELM} = \frac{1}{n} \sum_{i=1}^n (T_{PSOELMi} - Y_{PSOELMi})^2 \quad (16)$$

where i is the i_{th} particle; ω is a weight; $v_i(t)$ is the speed at time t ; $p_i(t)$ is the personal best position; $g_i(t)$ is the global best position; $x_i(t)$ is the position of the particle; c_1 and c_2 is a constant, which adjusting particle step size; r_1 and r_2 is a random number between (0,1); n is the number of samples; $T_{PSOELMi}$ is the network training output value; $Y_{PSOELMi}$ is the expected output value; and F_{PSOELM} is the fitness function defined by MSE.

The PSO algorithm is used to optimize the weight, which can be expressed as personal best position in the PSO, and threshold, which can be expressed as global best position in the PSO, of ELM. It consists of three steps (Assareh *et al.* 2010): (i) randomly generate initial position and initial velocity of all particles in the whole search space; (ii) update the velocity and position by fitness function; and (iii) calculate the value of all particles fitness and compare the current values with old personal best position and old global best position.

3.4 FM approach

Finite mixture (FM) model, a nonlinear regression equation, can be expressed by

$$P_{FMn}(t) = \sum_{i=1}^3 \omega_i P_{PMin}(t) \quad (17)$$

$$e_{FMn}(t) = P_n(t) - P_{FMn}(t) \quad (18)$$

$$\min Q = \sum_{n=1}^N (e_{FMn}(t))^2 \quad (19)$$

where $P_{PMin}(t)$ is the predicted output of the i_{th} prediction method at time t (P_{PM1n} can be represented for GRNN prediction, P_{PM2n} can be represented for BPNN prediction and P_{PM3n} can be represented for ELM prediction); $P_n(t)$ is the expected output; $P_{FMn}(t)$ is the predicted output based on the FM method; ω_i is the i_{th} weight factor, $\sum \omega_i = 1$ and $0 \leq \omega_i \leq 1$. $e_{FMn}(t)$ is the error between the predicted output and the expected output; and N is the number of data.

The Lagrange method, which is effective for solving the conditional extremum problem, is used to solve the parameter ω_i , expressed by

$$Q_{FM} = \sum_{n=1}^N (P_n(t) - \sum_{i=1}^3 \omega_i P_{PMin}(t))^2 + \lambda (\sum_{i=1}^3 \omega_i - 1) \quad (20)$$

$$\frac{\partial Q_{FM}}{\partial \omega_i} = \lambda - 2 \sum_{n=1}^N P_{PMin}(t) (P_n(t) - \sum_{i=1}^3 \omega_i P_{PMin}(t)) = 0 \quad (21)$$

$$\frac{\partial Q_{FM}}{\partial \lambda} = \sum_{i=1}^3 \omega_i - 1 = 0 \quad (22)$$

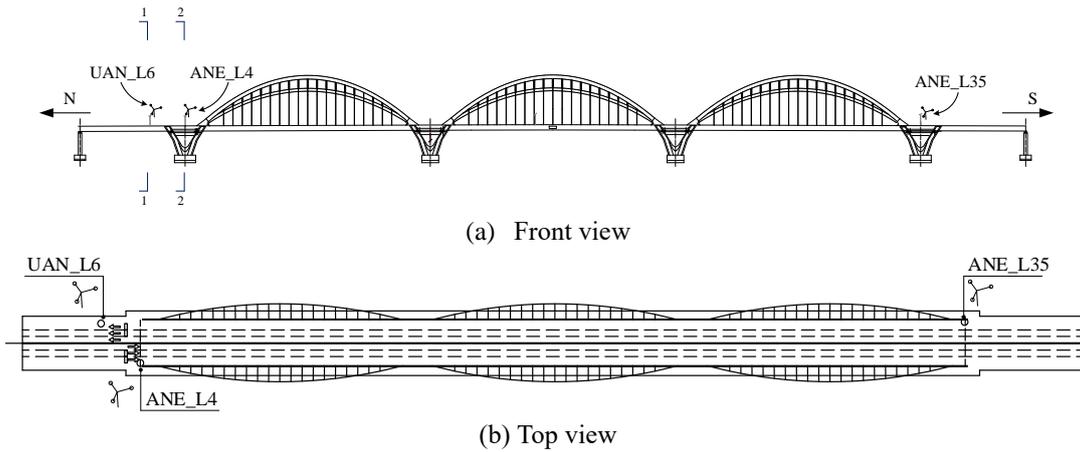


Fig. 1 Wind monitoring system installed on the investigated bridge

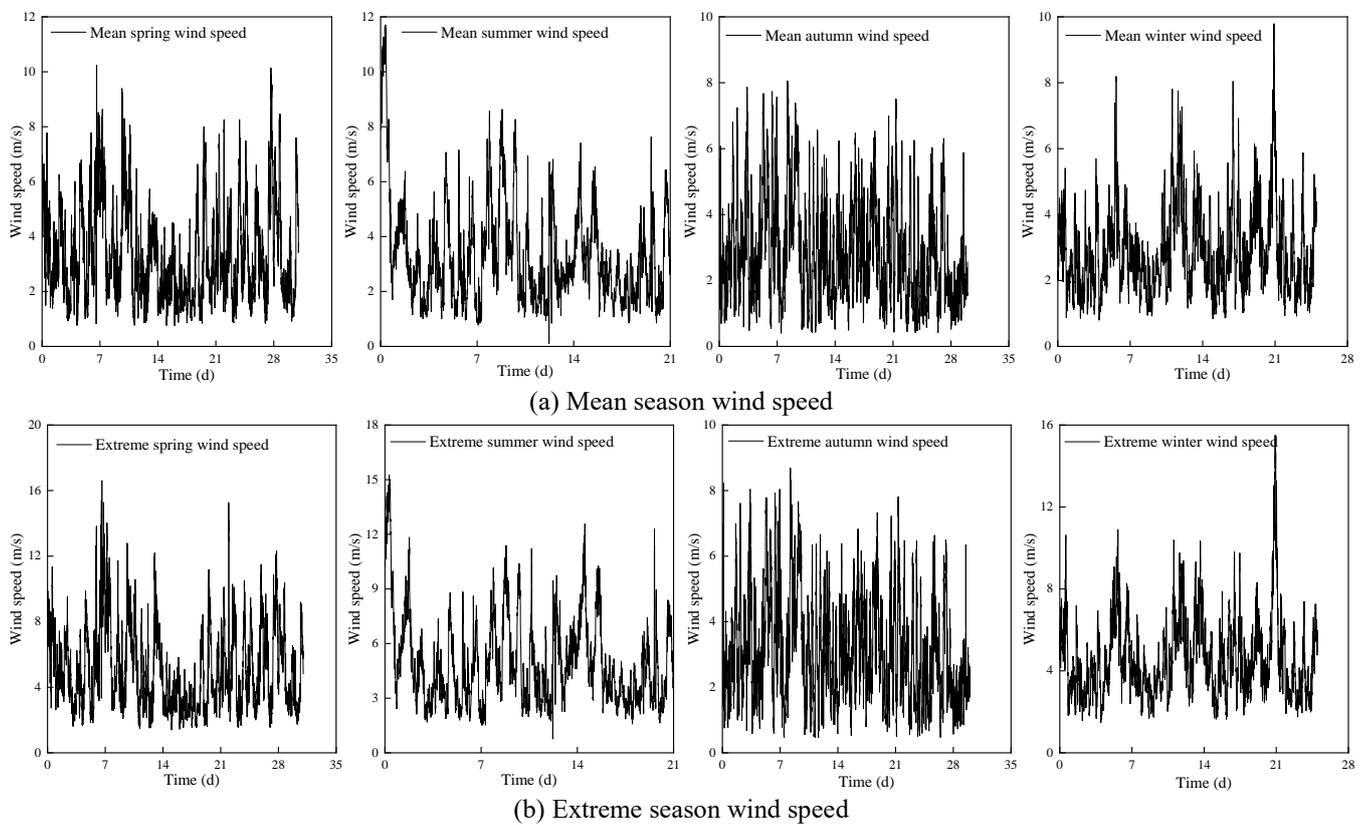


Fig. 2 Mean season wind speed and extreme season wind speed

where λ represents the Lagrange multiplier; and Q_{FM} represents the Lagrange function.

3.5 Model assessment

To evaluate the performance of the constructed prediction models, the mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE) are adopted as measure criteria.

The mean absolute error (MAE) of a model with respect to a test set is the mean of the absolute values of the

individual prediction errors in the test set. Each prediction error is the difference between the true and the predicted values, expressed by

$$MAE = \frac{1}{N} \sum_{t=1}^N |P_{EOt} - P_{POt}| \quad (23)$$

where N is the number of predictions; P_{EOt} is the value of expected output; and P_{POt} is the value of predicted output.

The mean absolute percentage error (MAPE) is a measure of prediction accuracy of a forecasting method in

statistics. It is defined by

$$MAPE = \frac{1}{N} \sum_{t=1}^N \frac{|P_{EOt} - P_{POt}|}{P_{EOt}} \quad (24)$$

The root mean square error (RMSE) is a frequently used measure of the difference between values predicted by a model and the values observed, expressed by

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N (P_{EOt} - P_{POt})^2} \quad (25)$$

4. Investigated bridge and wind field measurements

The Jiubao Bridge is located in Hangzhou, China. The total length of the bridge (including approach bridge) is 1855 m, and its main bridge is 3×210 m. One ultrasonic anemometer and two mechanical anemometers were installed at the 6 meters level above the bridge deck, as shown in Fig. 1. The sample frequencies of the ultrasonic anemometer and mechanical anemometer are 4 Hz and 0.1 Hz, respectively. The wind speed range of the ultrasonic anemometer is [0, 60] m/s with a resolution of 0.01 m/s, and its wind direction range is [0°, 360°] within an error of 0.1°. The wind speed measurement range of the mechanical anemometer is [0, 45] m/s, and its wind direction range is [0°, 360°] within an error of 0.1°.

The wind data measured by one ultrasonic anemometer from September 2014 to August 2015 are utilized to analyze the wind field characteristics. The wind data is categorized into four seasons: spring wind data (10-min mean spring wind speed and 10-min extreme spring wind speed), summer wind data (10-min mean summer wind speed and 10-min mean extreme summer wind speed), autumn wind data (10-min mean autumn wind speed and 10-min mean extreme autumn wind speed), and winter wind data (10-min mean winter wind speed and 10-min mean extreme winter

For the format requirements of the GRNN, BPNN and ELM input data, the wind speed needs to be preprocessed and classified. It follows three steps: (i) the wind speed data are arranged in a chronological order to establish a time series model; (ii) a small number of the first wind speed points express the current wind speed label; and (iii) the machine learning models are trained and then the wind speed data can be forecasted by the constructed models.

5. Wind speed forecasting by machine learning approaches

5.1 Spring wind speed forecasting

The mean spring wind speeds are forecasted by GRNN, BPNN, ELM, CV-GRNN, GA-BPNN, PSO-ELM and FM method. The results are shown in Fig. 3.

It can be seen that the prediction accuracy of ELM is better than GRNN and BPNN. Specifically, the RMSEs of ELM, GRNN, BPNN are 0.1755, 0.6428, 0.2567,

respectively, as shown in Fig. 4. As listed in Table 1, the weights of ELM, GRNN, BPNN are 1, 0, 0 respectively. Obviously, CV-GRNN, GA-BPNN, PSO-ELM and FM method can effectively improve the forecasting performance. For the CV-GRNN, its MAE, MAPE and RMSE are 60.94%, 69.94%, 53.00% less than those of GRNN. For the GA-BPNN, its MAE, MAPE and RMSE are 33.90%, 48.80%, 3.39% less than those of BPNN. The MAE and RMSE of PSO-ELM are 2.11%, 5.70% less than those of ELM, while the MAPE of PSO-ELM is 3.38% more than ELM.

For the FM method, its MAE, MAPE and RMSE are 77.76%, 82.45%, 72.70% less than those of GRNN. MAE, MAPE and RMSE of FM are 34.36%, 46.00%, 31.63% less than those of BPNN. It can be concluded that the performance of FM method depends on the components and is better than single model method.

The extreme spring wind speed is forecasted by GRNN, BPNN, ELM, CV-GRNN, GA-BPNN, PSO-ELM and FM method. The results are shown in Fig. 5.

It can be seen that the prediction accuracy of ELM is better than GRNN and BPNN. Specifically, the RMSEs of ELM, GRNN, BPNN are 0.2124, 0.6175, 0.2499, respectively, as shown in Fig. 6.

Table 1 Weights in FM method (for mean wind speed)

Weight	GRNN	BPNN	ELM
Spring wind	0	0	1
Summer wind	0	0.4892	0.5108
Autumn wind	0	0.04332	0.5668
Winter wind	0	0	1

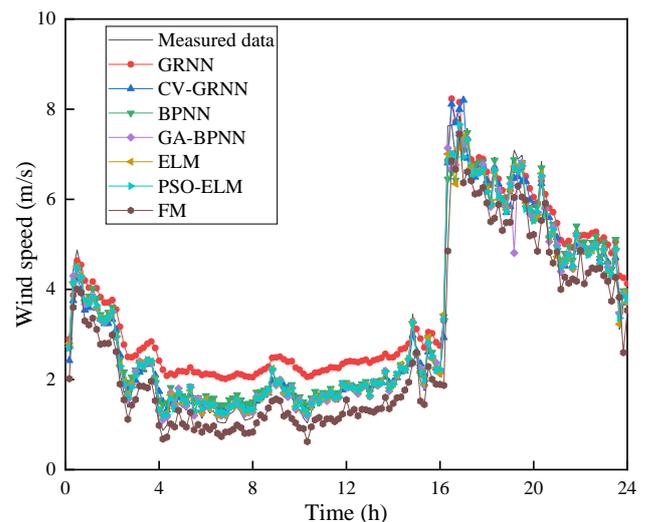


Fig. 3 Forecasting of mean spring wind speed by machine learning approaches

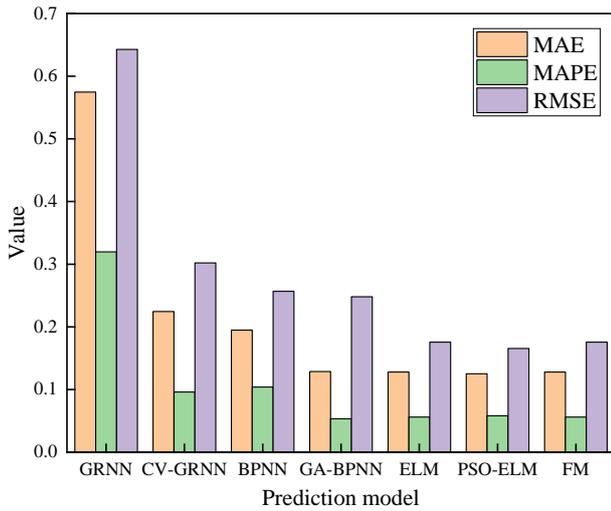


Fig. 4 Performance of machine learning approaches in forecasting mean spring wind speed

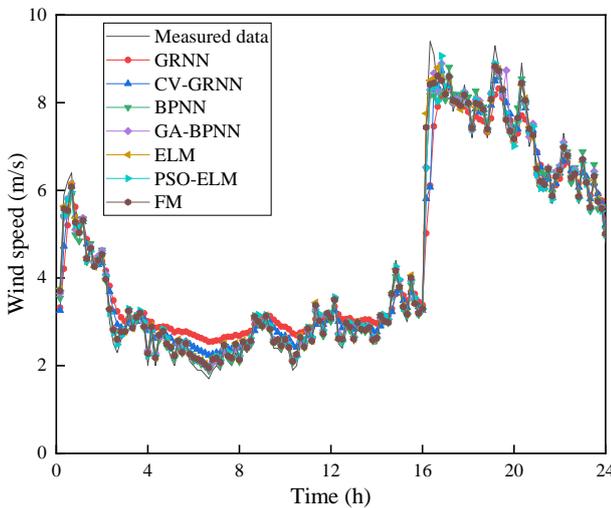


Fig. 5 Forecasting of extreme spring wind speed by machine learning approaches

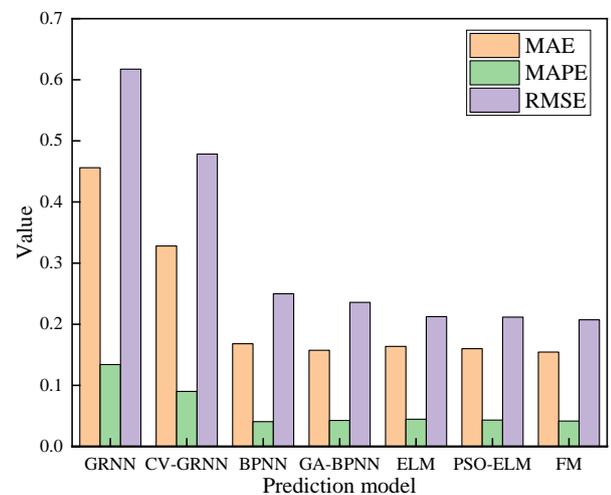


Fig. 6 Performance of machine learning approaches in forecasting extreme spring wind speed

Table 2 Weights in FM method (for extreme wind speed)

Weight	GRNN	BPNN	ELM
Extreme spring wind	0	0.2516	0.7484
Extreme summer wind	0	0.2178	0.7822
Extreme autumn wind	0	0.3382	0.6618
Extreme winter wind	0	0.3657	0.6343

As listed in Table 2, the weights of ELM, GRNN, BPNN are 0.7484, 0, 0.2516 respectively. Obviously, CV-GRNN, GA-BPNN, PSO-ELM and FM method can effectively improve the forecasting performance. For the GA-BPNN, its MAE, and RMSE are 6.42%, 5.64% less than those of BPNN, while the MAPE of GA-BPNN is 4.93% more than BPNN. For the GA-BPNN, its MAE, and RMSE are 6.42%, 5.64% less than those of BPNN, while the MAPE of GA-BPNN is 4.93% more than BPNN. The MAE and RMSE of PSO-ELM are 2.11%, 5.70% less than those of ELM, while the MAPE of PSO-ELM is 3.38% more than ELM.

For the FM method, its MAE, MAPE and RMSE are 66.14%, 68.98%, 66.45% less than those of GRNN. MAE and RMSE of FM are 8.15%, 17.09% less than those of BPNN, while the MAPE of FM is 2.46% more than BPNN. MAE, MAPE and RMSE of FM are 5.68%, 6.52%, 2.45% less than those of ELM. It can be concluded that the performance of FM method depends on the components and is better than single model method.

5.2 Summer wind speed forecasting

The mean summer wind speeds are forecasted by GRNN, BPNN, ELM, CV-GRNN, GA-BPNN, PSO-ELM and FM method are shown in Fig. 7.

As seen from Fig. 8, for the CV-GRNN, its MAE, MAPE and RMSE are 15.06%, 14.26%, 14.59% less than those of GRNN. For the GA-BPNN, its MAE, MAPE and RMSE are 5.54%, 11.06%, 4.90% less than those of BPNN. The MAE, MAPE and RMSE of PSO-ELM are 7.31%, 1.00%, 7.18% less than those of ELM.

For the FM method, its MAE, MAPE and RMSE are 66.53%, 68.19%, 67.60% less than those of GRNN. MAE, MAPE and RMSE of FM are 1.20%, 5.87%, 3.02% less than those of BPNN. MAE and RMSE of FM are 0.83%, 2.80% less than those of ELM, while the MAPE of FM is 4.25% more than ELM.

The extreme summer wind speeds are forecasted by GRNN, BPNN, ELM, CV-GRNN, GA-BPNN, PSO-ELM and FM method are shown in Fig. 9.

As seen from Fig. 10, for the CV-GRNN, its RMSE is 1.40% less than those of GRNN, while the MAE and MAPE of CV-GRNN are 9.04%, 13.06% more than those of GRNN. For the GA-BPNN, its MAE, MAPE and RMSE are 5.57%, 7.80%, 5.12% less than those of BPNN. The RMSE of PSO-ELM is 6.38% less than ELM, while the MAE and MAPE of PSO-ELM are 2.40%, 18.95% more than those of ELM.

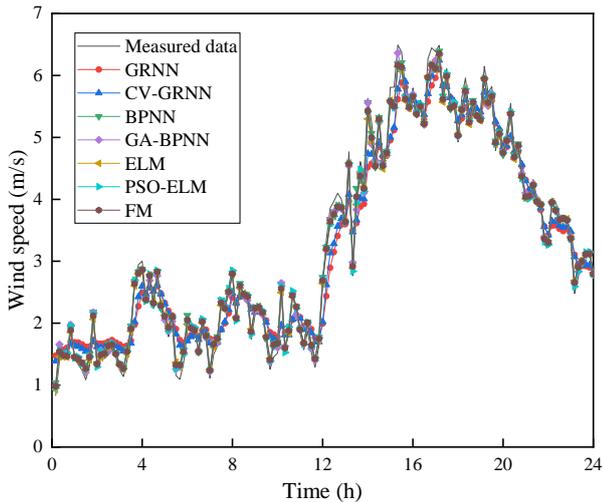


Fig. 7 Forecasting of mean summer wind speed by machine learning approaches

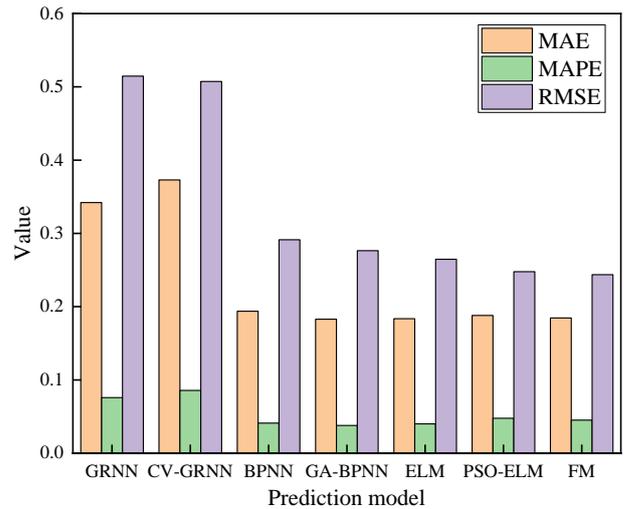


Fig. 10 Performance of machine learning approaches in forecasting extreme summer wind speed

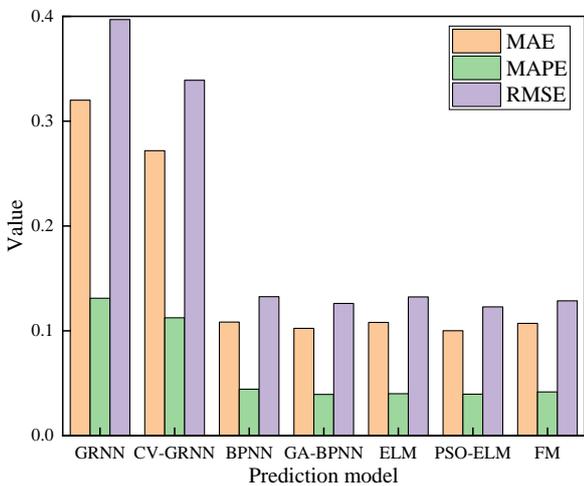


Fig. 8 Performance of machine learning approaches in forecasting mean summer wind speed

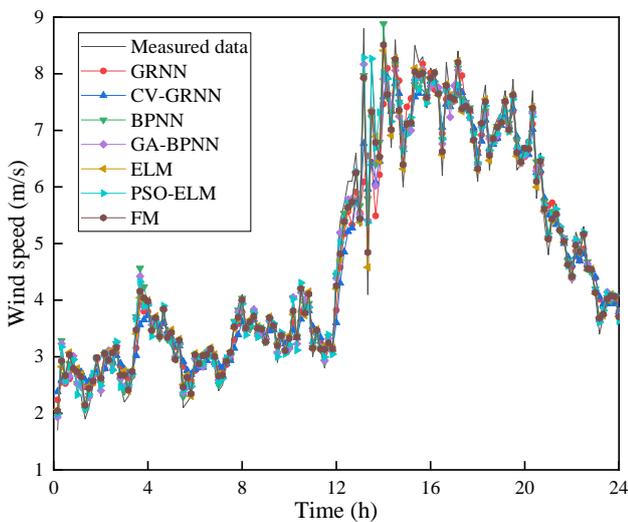


Fig. 9 Forecasting of extreme summer wind speed by machine learning approaches

For the FM method, its MAE, MAPE and RMSE are 45.99%, 40.50%, 52.64% less than those of GRNN. MAE and MAPE of FM are 4.70%, 16.34% less than those of BPNN, while the RMSE of FM is 10.00% more than BPNN. RMSE of FM is 7.93% less than ELM, while the MAE and MAPE of FM are 0.60%, 12.47% more than those of ELM.

5.3 Autumn wind speed forecasting

The mean autumn wind speeds are predicted by GRNN, BPNN, ELM, CV-GRNN, GA-BPNN, PSO-ELM and FM method are shown in Fig. 11.

As seen from Fig. 12, for the CV-GRNN, its MAE, MAPE and RMSE are 14.91%, 14.68%, 10.54% less than those of GRNN. For the GA-BPNN, MAE, MAPE and RMSE are 1.35%, 3.35%, 5.69% less than those of BPNN. The MAE, MAPE and RMSE of PSO-ELM are 13.08%, 10.60%, 22.16% less than those of ELM.

For the FM method, its MAE, MAPE and RMSE are 63.20%, 61.19%, 59.04% less than those of GRNN. MAE and MAPE of FM are 0.56%, 1.47% more than those of BPNN, while the RMSE of FM is 7.28% less than BPNN.

The extreme autumn wind speeds are predicted by GRNN, BPNN, ELM, CV-GRNN, GA-BPNN, PSO-ELM and FM method are shown in Fig. 13.

As seen from Fig. 14, for the CV-GRNN, its MAE, MAPE and RMSE are 19.47%, 14.65%, 26.65% less than those of GRNN. For the GA-BPNN, MAE, MAPE and RMSE are 7.88%, 7.26%, 7.83% less than those of BPNN. The MAE, MAPE and RMSE of PSO-ELM are 6.77%, 5.45%, 4.52% less than those of ELM.

For the FM method, its MAE, MAPE and RMSE are 56.01%, 53.62%, 56.94% less than those of GRNN. MAE, MAPE and RMSE of FM are 8.95%, 9.20%, 7.30% less than those of BPNN. MAE, MAPE and RMSE of FM are 1.91%, 0.19%, 2.08% less than those of ELM.

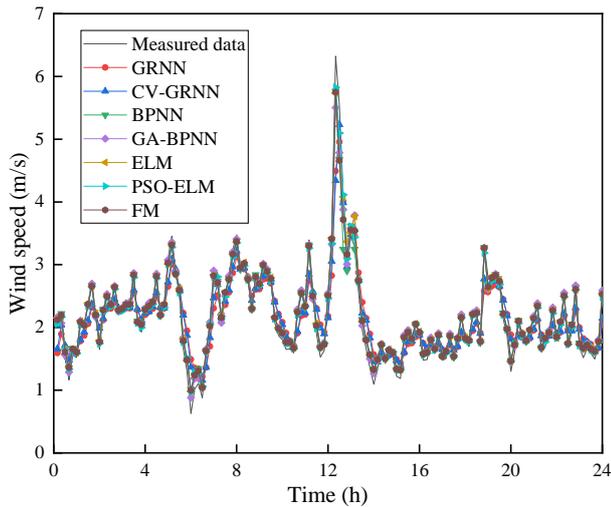


Fig. 11 Forecasting of mean autumn wind speed by machine learning approaches

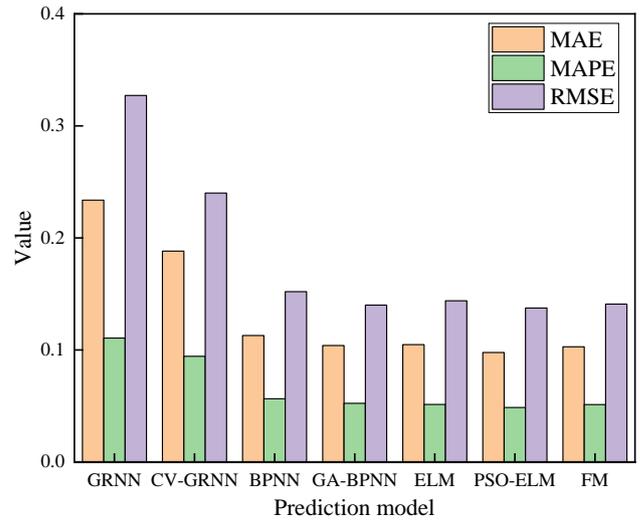


Fig. 14 Performance of machine learning approaches in forecasting extreme autumn wind speed

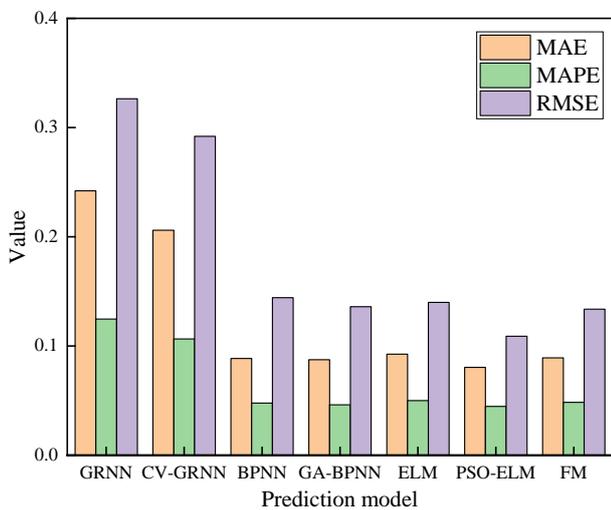


Fig. 12 Performance of machine learning approaches in forecasting mean autumn wind speed

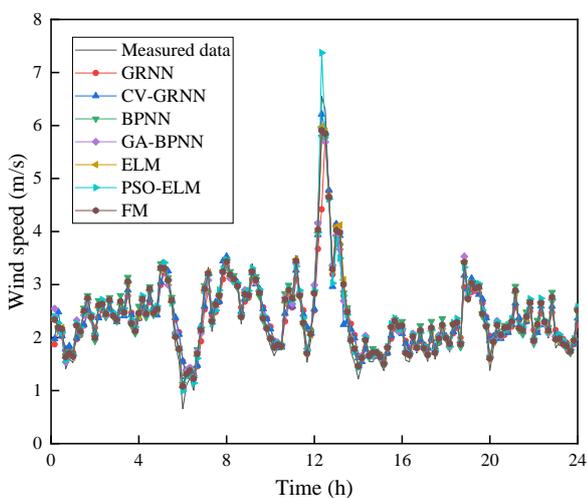


Fig. 13 Forecasting of extreme autumn wind speed by machine learning approaches

5.4 Winter wind speed forecasting

The mean winter wind speeds are predicted by GRNN, BPNN, ELM, CV-GRNN, GA-BPNN, PSO-ELM and FM method are shown in Fig. 15.

As can be seen from Fig. 16, for the CV-GRNN, its MAE, MAPE and RMSE are 5.70%, 13.44%, 4.84% less than those of GRNN. For the GA-BPNN, its MAE, MAPE and RMSE are 21.29%, 32.46%, 19.29% less than those of BPNN. The MAE, MAPE and RMSE of PSO-ELM are 16.65%, 17.62%, 19.58% less than those of ELM.

For the FM method, its MAE, MAPE and RMSE are 60.29%, 57.96%, 58.48% less than those of GRNN. MAE, MAPE and RMSE of FM are 30.06%, 33.22%, 25.62% less than those of BPNN.

The extreme winter wind speeds are predicted by GRNN, BPNN, ELM, CV-GRNN, GA-BPNN, PSO-ELM and FM method are shown in Fig. 17.

As can be seen from Fig. 18, for the CV-GRNN, its MAE, MAPE and RMSE are 15.55%, 20.26%, 15.13% less than those of GRNN. For the GA-BPNN, its MAE, MAPE and RMSE are 4.72%, 5.18%, 5.61% less than those of BPNN. The MAE, MAPE and RMSE of PSO-ELM are 2.57%, 5.67%, 1.04% less than those of ELM.

For the FM method, its MAE, MAPE and RMSE are 58.25%, 60.22%, 56.87% less than those of GRNN. MAE, MAPE and RMSE of FM are 1.24%, 3.60%, 1.66% less than those of BPNN. MAE and RMSE of FM are 0.25%, 0.54% less than those of ELM, while the MAPE is 1.18% more than ELM.

6. Conclusions

In this paper, various machine learning methods are explored to forecast the wind speed. A total of seven machine learning methods are under consideration, namely

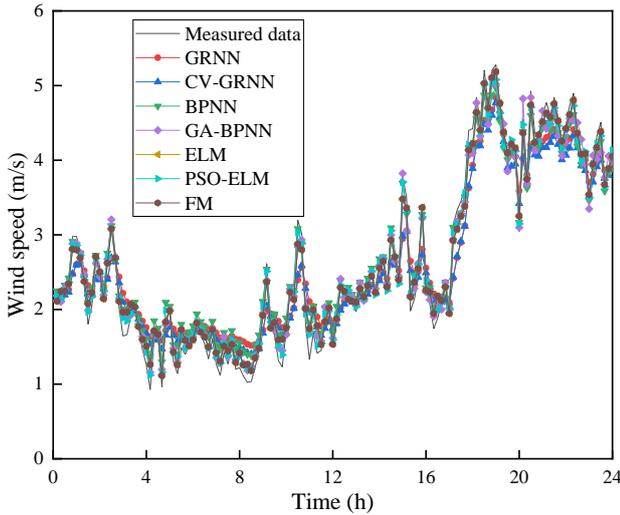


Fig. 15 Forecasting of mean winter wind speed by machine learning approaches

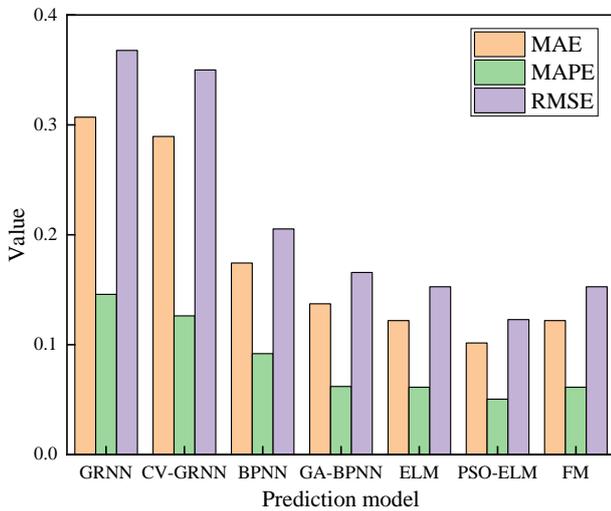


Fig. 16 Performance of machine learning approaches in forecasting mean winter wind speed

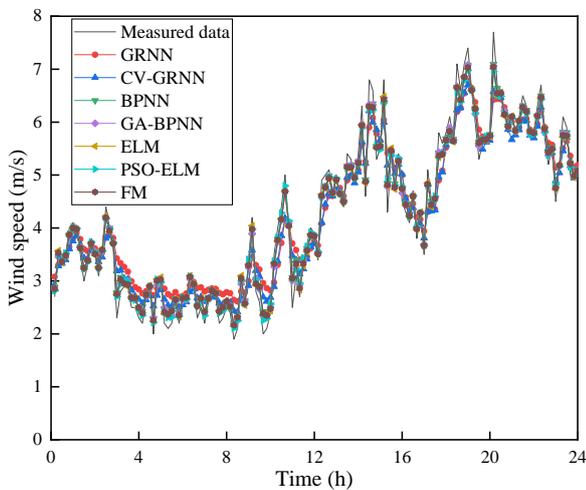


Fig. 17 Forecasting of extreme winter wind speed by machine learning approaches

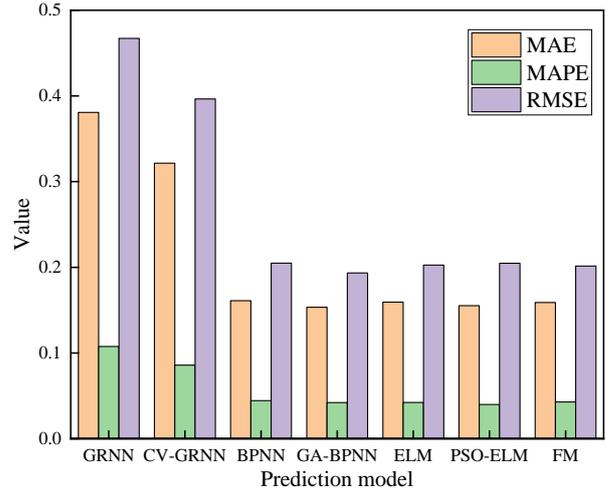


Fig. 18 Performance of machine learning approaches in forecasting extreme winter wind speed

GRNN, BPNN, ELM, CV-GRNN, GA-BPNN, PSO-ELM, and FM. In particular, CV-GRNN, GA-BPNN, PSO-ELM belong to the optimization algorithm-assisted machine learning approaches, and FM is a hybrid machine learning approach consisting of GRNN, BPNN, and ELM.

One-year wind speed monitoring data collected by an SHM system installed on the Jiubao Bridge is adopted to demonstrate the applicability of these machine learning approaches in forecasting the wind speed. The forecasting performance of these machine learning methods is fully compared. The main conclusions drawn from this study are summarized as follows:

1. Overall, the optimization algorithm-assisted and hybrid machine learning approaches have better forecasting performance than the traditional machine learning methods (GRNN, BPNN, and ELM).
2. For the traditional machine learning methods, the forecasting accuracy of ELM is better than GRNN and BPNN. Optimization algorithms integrated into the machine learning methods can improve the forecasting performance. For the optimization algorithm-assisted machine learning methods, the forecasting performance of PSO-ELM is better than CV-GRNN and GA-BPNN.
3. The forecasting accuracy of FM method largely depends on the machine learning methods and is better than machine learning methods, which are components of FM method. The hybrid machine learning approach (e.g., FM) and the optimization algorithm-assisted ones have similar forecasting performance.

Acknowledgments

The work described in this paper was jointly supported by the National Science Foundation of China (Grant Nos. 51822810, 51878235, and 51778574), the Zhejiang Provincial Natural Science Foundation of China (Grant No. LR19E080002), and the Fundamental Research Funds for the Central Universities of China (Grant No. 2019XZZX004-01).

References

- Assareh, E., Behrang, M.A., Assari, M.R. and Ghanbarzadeh, A. (2010), "Application of pso (particle swarm optimization) and ga (genetic algorithm) techniques on demand estimation of oil in iran", *Energ.*, **35**(12), 5223-5229. DOI: 10.1016/j.energy.2010.07.043.
- Beheshti, Z., and Shamsuddin, S.M.H. (2014), "Capso: centripetal accelerated particle swarm optimization", *Inf. Sci.*, **258**, 54-79. DOI: 10.1016/j.ins.2013.08.015.
- Braganeto, U.M. and Dougherty, E.R. (2004), "Is cross-validation valid for small-sample microarray classification?", *Bioinform.*, **20**(3), 374-380. DOI: 10.1093/bioinformatics/btg419.
- Chen, K. and Yu, J. (2014), "Short-term wind speed prediction using an unscented kalman filter based state-space support vector regression approach", *Appl. Energ.*, **113**, 690-705. DOI: 10.1016/j.apenergy.2013.08.025.
- Cheung, K.S., Langevin, A. and Delmaire, H. (1997), "Coupling genetic algorithm with a grid search method to solve mixed integer nonlinear programming problems", *Comput. Math. Appl.*, **34**(12), 13-23. DOI: 10.1016/s0898-1221(97)00229-0.
- Cigizoglu, H.K. (2005), "Generalized regression neural network in monthly flow forecasting", *Civ. Eng. Environ. Syst.*, **22**(2), 71-81. DOI: 10.1080/10286600500126256.
- Diana, G. and Tommasi, C. (2002), "Cross-validation methods in principal component analysis: a comparison", *Stat. Method. Appl.*, **11**(1), 71-82. DOI: 10.1007/bf02511446.
- Ding, S., Su, C. and Yu, J. (2011), "An optimizing bp neural network algorithm based on genetic algorithm", *Artif. Intell. Rev.*, **36**(2), 153-162. DOI: 10.1007/s10462-011-9208-z.
- Engelbrecht, A.P. (2006), "Fundamentals of computational swarm intelligence", Wiley, Hoboken, NJ, USA.
- Guo, Z., Wu, J., Lu, H. and Wang, J. (2011), "A case study on a hybrid wind speed forecasting method using BP neural network", *Knowledge-Based Syst.*, **24**(7), 1048-1056. DOI: 10.1016/j.knsys.2011.04.019.
- Han, F., Yao, H.F. and Ling, Q.H. (2013), "An improved evolutionary extreme learning machine based on particle swarm optimization", *Neurocomputing*, **116**, 87-93. DOI: 10.1016/j.neucom.2011.12.062.
- Heimes, F. and van Heuveln, B. (1998), "The normalized radial basis function neural network", *P. IEEE. Int. Conf.*, **2**, 1609-1614. DOI: 10.1109/ICSMC.1998.728118.
- Hocaoglu, F.O. and Kurban, M. (2007), "The effect of missing wind speed data on wind power estimation", *Int. Conf. Intell. Data Eng. Autom. Lea.*, 107-114. DOI: 10.1007/978-3-540-77226-2_12.
- Holland, J.H. (1973), "Genetic algorithms and the optimal allocation of trials", *Siamj. Comput.*, **2**(2), 88-105. DOI: 10.1137/0202009.
- Huang, D.M., He, S.Q., He, X.H. and Zhu, X. (2017), "Prediction of wind loads on high-rise building using a bp neural network combined with pod", *J. Wind Eng. Ind. Aerod.*, **170**, 1-17. DOI: 10.1016/j.jweia.2017.07.021.
- Huang, G.B., Zhu, Q.Y. and Siew, C.K. (2006), "Extreme learning machine: theory and applications", *Neurocomputing*, **70**(1-3), 489-501. DOI: 10.1016/j.neucom.2005.12.126.
- Huang, G.B. and Lei, C. (2007), "Convex incremental extreme learning machine", *Neurocomputing*, **70**(16), 3056-3062. DOI: 10.1016/j.neucom.2007.02.009.
- Huang, K., Dai, L. and Huang, S. (2010), "Wind prediction based on improved bp artificial neural network in wind farm", *Int. Conf. El. Control Eng.*, 2548-2551. DOI: 10.1109/iCECE.2010.630.
- Jadid, M.N. and Fairbairn, D.R. (1994), "The application of neural network techniques to structural analysis by implementing an adaptive finite-element mesh generation", *Artif. Intel. Eng. Des. Anal. Manuf.*, **8**(3), 177-191. DOI: 10.1017/S0890060400001979.
- Jiang, P. and Chen, J. (2016), "Displacement prediction of landslide based on generalized regression neural networks with k-fold cross-validation", *Neurocomputing*, **198**(7), 40-47. DOI: 10.1016/j.neucom.2015.08.118.
- Kai, C., Qi, L., Yao, L. and Yong, D. (2016), "Robust regularized extreme learning machine for regression using iteratively reweighted least squares", *Neurocomputing.*, **230**, 345-358. DOI: 10.1016/j.neucom.2016.12.029.
- Kassa, Y., Zhang, J., Zheng, D. and Wei, D. (2016), "A ga-bp hybrid algorithm based ann model for wind power prediction", *IEEE. Smart Energ. Grid Eng.*, 158-163. DOI: 10.1109/SEGE.2016.7589518.
- Kennedy, J. and Eberhart, R. (1995), "Particle swarm optimization", *P. IEEE. Int. Conf. Neur. Net.*, **4**, 1942-1948. DOI: 10.1109/ICNN.1995.488968.
- Khalid, M. and Savkin, A.V. (2012), "A method for short-term wind power prediction with multiple observation points", *IEEE T. Power Syst.*, **27**(2), 579-586. DOI: 10.1109/tpwrs.2011.2160295.
- Kumar, G. and Malik, H. (2016), "Generalized regression neural network based wind speed prediction model for western region of india", *P. Comput. Sci.*, **93**, 26-32. DOI: 10.1016/j.procs.2016.07.177.
- Landberg, L. (1999), "Short-term prediction of the power production from wind farms", *J. Wind Eng. Ind. Aerod.*, **80**(1-2), 207-220. DOI: 10.1016/S0167-6105(98)00192-5.
- Li, X., Liu, Y. and Xin, W. (2009), "Wind speed prediction based on genetic neural network", *IEEE. Conf. Ind. El. Appl.*, 2448-2451. DOI: 10.1109/ICIEA.2009.5138642.
- Lei, M., Shiyang, L., Chuanwen, J., Hongling, L. and Yan, Z. (2009), "A review on the forecasting of wind speed and generated power", *Renew. Sust. Energ. Rev.*, **13**(4), 915-920. DOI: 10.1016/j.rser.2008.02.002.
- Lee, C.Y. and He, Y.L. (2012), "Wind prediction based on general regression neural network", *Int. Conf. Intell. Syst. Des. Eng. Appl.*, 617-620. DOI: 10.1109/ISdea.2012.520.
- Li, H.Z., Guo, S., Li, C.J. and Sun, J.Q. (2013), "A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm", *Knowledge-Based Syst.*, **37**(2), 378-387. DOI: 10.1016/j.knsys.2012.08.015.
- Lazarevska, E. (2016), "Wind speed prediction with extreme learning machine", *IEEE. Int. Conf. Intell. Syst.*, 154-159. DOI: 10.1109/IS.2016.7737415.
- Liu, H., Mi, X. and Li, Y. (2018), "An experimental investigation of three new hybrid wind speed forecasting models using multi-decomposing strategy and elm algorithm", *Renew. Energ.*, **123**, 694-705. DOI: doi.org/10.1016/j.renene.2018.02.092.
- Maulik, U. and Bandyopadhyay, S. (2000), "Genetic algorithm-based clustering technique", *Pattern Recognit.*, **33**(9), 1455-1465. DOI: 10.1016/S0031-3203(99)00137-5.
- McLachlan, G.J. and Peel, D. (2000), "Finite mixture models", Wiley, New York, N.Y., USA.
- Ni, Y.Q., Ye, X.W. and Ko, J.M. (2010), "Monitoring-based fatigue reliability assessment of steel bridges: analytical model and application", *J. Eng. Mech.*, **136**(12), 1563-1573. DOI: 10.1061/(ASCE)ST.1943-541X.0000250.
- Ni, Y.Q., Ye, X.W. and Ko, J.M. (2012), "Modeling of stress spectrum using long-term monitoring data and finite mixture distributions", *J. Eng. Mech.*, **138**(2), 175-183. DOI: 10.1061/(ASCE)EM.1943-7889.0000313.
- Ping, J., Zeng, Z., Chen, J. and Huang, T. (2014), "Generalized regression neural networks with k-fold cross-validation for displacement of landslide forecasting", *Adv. Neur. Net.*, 533-541. DOI: 10.1007/978-3-319-12436-0_59.

- Refaeilzadeh, P., Lei, T. and Liu, H. (2016), "Cross-validation", *Encyclopedia of Database Sys.*, 532-538. DOI: 10.1007/978-0-387-39940-9_565.
- Shao, Z., Meng, J.E. and Ning, W. (2016), "An efficient leave-one-out cross-validation-based extreme learning machine (eloolm) with minimal user intervention", *IEEE T. Cybern.*, **46**(8), 1939-1951. DOI: 10.1109/TCYB.2015.2458177.
- Specht, D.F. (1991), "A general regression neural network", *IEEE T. Neur. Net.*, **2**(6), 568-576. DOI: 10.1109/72.97934.
- Specht, D.F. (1992), "Enhancements to probabilistic neural networks", *Neur. Net.*, 761-768. DOI: 10.1109/IJCNN.1992.287095.
- Specht, D.F. (1993), "The general regression neural network-rediscovered", *Neur. Net.*, **6**(7), 1033-1034. DOI: 10.1016/S0893-6080(09)80013-0.
- Werbos, P.J. (1974), "Beyond regression: new tools for prediction and analysis in the behavioral sciences" *Harvard University*, Cambridge, MA.
- Wen, X.L., Huang, J.C., Sheng, D.H. and Wang, F.L. (2010), "Conicity and cylindricity error evaluation using particle swarm optimization" *Precis. Eng.*, **34**(2), 338-344. DOI: 10.1016/j.precisioneng.2009.08.002.
- Wang, S., Zhang, N., Wu, L. and Wang, Y. (2016), "Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and ga-bp neural network method." *Renew. Energ.*, **94**, 629-636. DOI: 10.1016/j.renene.2016.03.103.
- Wang, Y. and Peng, H. (2018), "Underwater acoustic source localization using generalized regression neural network." *J. Acoust. Soc. Am.*, **143**(4), 2321-2331. DOI: 10.1121/1.5032311.
- Xu, R.L., Xu, X., Zhu, B. and Chen, M. (2011), "The application of genetic-neural network on wind power prediction" *Int. Conf. Inf. Comput. Appl.*, 379-386. DOI: 10.1007/978-3-642-27452-7_52.
- Yang, D. and Han, F. (2014), "An improved ensemble of extreme learning machine based on attractive and repulsive particle swarm optimization" *Int. Conf. Intell. Comput.*, 213-220. DOI: 10.1007/978-3-319-09333-8_23.
- Ye, X.W., Ni, Y.Q., Wong, K.Y. and Ko, J.M. (2012), "Statistical analysis of stress spectra for fatigue life assessment of steel bridges with structural health monitoring data", *Eng. Struct.*, **45**, 166-176. DOI: 10.1016/j.engstruct.2012.06.016.
- Ye, X.W., Ni, Y.Q., Wai, T.T., Wong, K.Y., Zhang, X.M. and Xu, F. (2013), "A vision-based system for dynamic displacement measurement of long-span bridges: algorithm and verification", *Smart. Struct. Syst.*, **12**(3), 363-379. https://doi.org/10.12989/sss.2013.12.3_4.363.
- Ye, X.W., Yi, T.H., Wen, C. and Su, Y.H. (2015a), "Reliability-based assessment of steel bridge deck using a mesh-insensitive structural stress method", *Smart. Struct. Syst.*, **16**(2), 367-382. <https://doi.org/10.12989/sss.2015.16.2.367>.
- Ye, X.W., Yi, T.H., Dong, C.Z., Liu, T. and Bai, H. (2015b), "Multi-point displacement monitoring of bridges using a vision-based approach", *Wind Struct.*, **20**(2), 315-326. <https://doi.org/10.12989/was.2015.20.2.315>.
- Ye, X.W., Dong, C.Z. and Liu, T. (2016a), "Force monitoring of steel cables using vision-based sensing technology: methodology and experimental verification", *Smart. Struct. Syst.*, **18**(3), 585-599. <https://doi.org/10.12989/sss.2016.18.3.585>.
- Ye, X.W., Yi, T.H., Dong, C.Z. and Liu, T. (2016b), "Vision-based structural displacement measurement: system performance evaluation and influence factor analysis", *Measurement*, **88**, 372-384. DOI: 10.1016/j.measurement.2016.01.024.
- Ye, X.W., Dong, C.Z. and Liu, T. (2016c), "Image-based structural dynamic displacement measurement using different multi-object tracking algorithms", *Smart. Struct. Syst.*, **17**(6), 935-956. <https://doi.org/10.12989/sss.2016.17.6.935>.
- Ye, X.W., Su, Y.H., Xi, P.S., Chen, B. and Han, J.P. (2016d), "Statistical analysis and probabilistic modeling of WIM monitoring data of an instrumented arch bridge", *Smart. Struct. Syst.*, **17**(6), 1087-1105. <https://doi.org/10.12989/sss.2016.17.6.1087>.
- Ye, X.W., Yi, T.H., Su, Y.H., Liu, T. and Chen, B. (2017), "Strain-based structural condition assessment of an instrumented arch bridge using FBG monitoring data", *Smart. Struct. Syst.*, **20**(2), 139-150. <https://doi.org/10.12989/sss.2017.20.2.139>.